



Generalised Manifolds as Basic Objects of General Relativity

Joanna Luc¹ 

Received: 2 January 2019 / Accepted: 14 August 2019 / Published online: 28 August 2019
© The Author(s) 2019

Abstract

In this paper non-Hausdorff manifolds as potential basic objects of General Relativity are investigated. One can distinguish four stages of identifying an appropriate mathematical structure to describe physical systems: kinematic, dynamical, physical reasonability, and empirical. The thesis of this paper is that in the context of General Relativity, non-Hausdorff manifolds pass the first two stages, as they enable one to define the basic notions of differential geometry needed to pose the problem of the evolution-distribution of matter and are not in conflict with the Einstein equations. With regard to the third stage, various potential conflicts with physical reasonability conditions are considered with a tentative conclusion that non-Hausdorff manifolds are more likely to pass this stage than is typically assumed. When dealing with some of these problems, the modal interpretation of non-Hausdorff manifolds is invoked, according to which they represent bundles of alternative possible spacetimes rather than single spacetimes.

Keywords Non-Hausdorff manifolds · Generalised manifolds · Indeterminism · Physical reasonability

1 Introduction

The topic of this paper is the mathematical properties of non-Hausdorff manifolds, especially those that seem to be physically relevant, and the possible use of these objects in General Relativity (GR). Most GR textbooks focus solely on manifolds that satisfy the Hausdorff condition (defined later); typically, the term ‘manifold’ is used as a synonym of ‘Hausdorff manifold’, and the class that encompasses both Hausdorff and non-Hausdorff manifolds is called the class of ‘generalised manifolds’. The main question of this paper is to what extent the standard exclusion of non-

✉ Joanna Luc
joanna.luc.poczta@gmail.com

¹ Institute of Philosophy, Jagiellonian University, Grodzka 52, 31-044 Kraków, Poland

Hausdorff manifolds from the scope of investigations in physics is justified, and if it is justified, what the basis and conceptual significance are for such an exclusion.

Some of the crucial differences between Hausdorff and non-Hausdorff manifolds are connected with the issue of determinism and indeterminism. Non-Hausdorff manifolds allow for two forms of indeterminism that are excluded when we impose the Hausdorff condition: at the level of geodesics and at the level of entire spacetimes. Therefore, it will be suggested that non-Hausdorff manifolds are better interpreted as representing bundles of alternative possible spacetimes rather than single spacetimes, which alleviates some of the problems associated with them.

The paper is structured as follows. In Sect. 2, the notion of a basic object of GR is introduced. In Sect. 3 a scheme is proposed for identifying an appropriate mathematical structure to describe physical systems, which consists of four stages: kinematic, dynamical, physical reasonability, and empirical. In Sect. 4, it is argued that in the case of GR, the necessary condition for passing a kinematic stage is the satisfaction of a functional criterion, according to which a basic object of GR should provide a mathematical framework for posing the problem of the evolution-distribution of matter. Section 5 provides a formal definition of the Hausdorff condition and includes definitions of the different types of manifolds (i.e. Hausdorff, non-Hausdorff, and generalised). In Sect. 6, the properties of non-Hausdorff manifolds are reviewed, which lays the groundwork for the discussion of objections against the use of non-Hausdorff manifolds in physics in Sect. 7. Section 8 answers these objections and concludes that non-Hausdorff manifolds pass the kinematic and dynamical stages and that their situation regarding the third stage is not as problematic as typically assumed. To deal with some of the objections, the modal interpretation of non-Hausdorff manifolds is introduced. Finally, Sect. 9 identifies two senses of modality that were involved in the considerations discussed in the previous sections: first in connection with indeterminism and then with the four stages. Section 10 provides a short summary.

2 The Notion of the Basic Object of GR

The basic object of GR is the main object this theory is about and with respect to which all other objects in the theory are defined. The expression ‘with respect to which all other objects in the theory are defined’ does not mean that the basic object of GR exhausts this theory’s content and that all other objects are defined *in terms of it*; it is the much more modest claim that all other objects somehow presume it and are defined *with respect to it* (or *on it*). Other objects that GR deals with are different types of matter, represented either by matter fields or by point particles, both of which are defined with respect to the basic object of GR (i.e. they are represented by functions of spatiotemporal variables).

The term ‘basic object’ should be understood ontologically, that is, as one of the objects to which the theory refers. Mathematically, the basic object of GR is represented by a differential manifold (perhaps with some additional structures, such as a metric). Therefore, the necessary condition for being the basic object of GR is being a physical counterpart of a differential manifold; however, this is likely to be not sufficient, as one can imagine using the formalism of GR to model phenomena quite

remote from those typical for GR. In Sect. 4, I will impose a further condition on basic objects of GR.

The notion of the basic object of GR (or an equivalent one) is not used in the literature,¹ and there is a reason for this—it is always assumed that it is clear what the basic object of GR is, namely: spacetime. However, the hypothesis defended here is that spacetime is not the only possible option for a basic object of GR. In what follows, it will be argued that if we define the basic object of GR by some minimal conditions it should possess, then it turns out that other objects which are more naturally interpreted as bundles of alternative possible spacetimes rather than single spacetimes belong to this category. Therefore, to allow for both options, the more general notion of a basic object of GR is needed.

3 Four Stages of Identifying the Appropriate Mathematical Structures to Represent the Physical World

The aim of theoretical physics is to find mathematical structures that adequately represent physical reality. This suggests the following two-stage picture: physicists first propose a mathematical structure, and then, via an empirical investigation, they decide whether the structure adequately represents the physical world or a part of it. However, this way of proceeding is not very practical, as the number of mathematical structures is too large to try them one by one. Instead, physicists start from some set of mathematical structures and then systematically narrow it down to get the desired result. This ‘narrowing down’ can be done on a purely empirical basis, but it can also have a mixed, empirical-theoretical character. The first case is typical for situations where one has a parametrised family of models; then, by fitting parameters to the available data, some ranges of these parameters can be excluded. In the next few paragraphs, a scheme for empirical-theoretical narrowing down is proposed, which differs from the purely empirical one in that each stage has some conceptual (and perhaps also a metaphysical, especially modal) significance. The scheme is as follows:

1. Kinematic stage: the choice of the basic set-up of a theory, including the space of possible states of its objects. At this stage, no constraints in the form of equations are imposed.
2. Dynamical stage: the choice of mathematical constraints in the form of dynamical equations that the objects should satisfy.

¹ In the context of quantum theory, there is a similar notion, namely ‘primitive ontology of a theory’, which is said to be “what the theory is fundamentally about” [1, p. 363]. However, a closer look reveals significant differences between the two notions, which are the consequence of differences in context and the purpose of introducing them. Primitive ontology consists of entities living *in* spacetime, which are fundamental building blocks of everything else [2, p. 60]; they should be distinguished from other objects used in a theory, whose role is only auxiliary, such as the wave function, which does not have physical significance of its own, but only as a tool in the description of the behaviour of the objects belonging to primitive ontology. The role of the notion of the basic object of GR is quite different: it is not intended to capture *all* physically significant objects that can be described by GR, and it is also not intended to distinguish between physically real and auxiliary objects in a theory (or even between more fundamental and less fundamental ones); rather, it is introduced to avoid presuming that spacetime is the only possible referent of the central mathematical structure of GR, that is, a differential manifold.

3. Physical reasonability stage: the choice of additional physical reasonability conditions, which should be satisfied by the objects.
4. Empirical stage: finding the model of a theory that is realised (or approximately realised) in the actual world.

This scheme is not supposed to be an original proposal. The distinction between kinematically and dynamically possible models is well known and popular in the philosophy of physics literature;² also, the importance of physical reasonability as providing additional constraints is recognised in the literature, especially in the context of GR (see, e.g., [5,6]), although such conditions may appear in other theories as well.

The stages differ from each other in the following way. The kinematic stage is distinguished by the absence of any constraints in the form of equations, which are introduced at the second and perhaps the third stage. The dynamical stage imposes dynamical equations (and in the relativistic case also constraints on the initial conditions associated with these dynamical equations). The physical reasonability stage encompasses various additional constraints on the behaviour of the solutions of these dynamical equations.

In the case of a basic object of GR, the above four stages can be specified in the following way:

1. Kinematic stage: differential geometry as a basic set-up; basic objects of GR are differential manifolds with metrics; matter is represented by an energy tensor.³
2. Dynamical stage: a basic object of GR should satisfy the Einstein equations.
3. Physical reasonability stage: a basic object of GR should satisfy some physical reasonability conditions, for example, causality conditions and energy conditions; the exact list of the conditions required is disputable.⁴
4. Empirical stage: addressed by observational astronomy, astrophysics and cosmology.

The hierarchy of stages does not need to correspond to temporal succession in the actual scientific work. All the stages are empirically motivated in some way, not just the last one. Kinematics, dynamics and physical reasonability principles should be chosen such that their defining conditions are not violated by the actual physical world. Furthermore, physicists often investigate theories that are not mathematically fully understood, so they may consider the equation of motion (i.e., the dynamical stage) without explicitly specifying all the details of the space of states (i.e., the kinematic stage).

The crucial fact is that in contrast to purely empirical narrowing down, each of the stages in the above scheme has its own individual conceptual significance. The elimination of certain ranges of parameters (in purely empirical narrowing down) concerns just numbers, and its consecutive phases do not have such a clear meaning.

² For exemplary uses of this distinction, see [3, p. 154] or [4, p. 702]. The related distinction between kinematics and dynamics is standard in physics.

³ Strictly speaking, an energy tensor does not represent all ‘aspects’ of matter, only its properties that are relevant for gravitational interactions; a full representation of matter is given by matter fields which possess energy and momentum represented by this tensor.

⁴ The notion of physical reasonable spacetime is used in the context of GR by physicists and philosophers of physics. For the list of potential physical reasonability conditions, see, for example, [5].

Therefore, the scheme should be read as proposing a kind of conceptual order, which can be identified in different theoretical contexts in physics.

4 Kinematic Stage in the Case of a Basic Object of GR

To perform the first stage, some initial heuristics are required. I argue elsewhere [7] that in the case of GR, as well as other spacetime theories, such a heuristic may be provided by functional considerations: a basic object of GR, which is typically just spacetime, can be defined by the function it is supposed to play with respect to other objects in the theory, which is providing the basic framework for posing the problem of the evolution-distribution of matter.

To elaborate on this issue, in [7], I explain in more detail how the proposed function should be understood and the benefits of this way of thinking about spacetime theories. Regarding the first issue, I use a non-standard notion of evolution-distribution where the hyphen indicates that this should be understood as a single idea whose two aspects, evolution and distribution, can be clearly distinguished only in some cases. ‘Distribution’ refers to the arrangement of matter in space, whereas ‘evolution’ refers to the evolution in time of such an arrangement. In pre-relativistic physics (and according to common sense), space and time are separate structures, so we can associate space with distribution and time with evolution. However, this is no longer the case in relativistic physics, so that is why the hyphenated term is used. The level of this inseparability varies. In SR and globally hyperbolic solutions of GR, this separation is still possible, but it can be done in many ways, none of which are distinguished. For solutions of GR that are not globally hyperbolic, space and time cannot be separated even in this weaker sense.⁵

The formulation of the problem of evolution-distribution of matter amounts to providing consistent equations prescribing the evolution-distribution of matter (in other words, the laws governing the evolution-distribution of matter) which have at least one solution. The term ‘formulation’ should be, therefore, understood in a rather broad manner in that it does not imply the well-posedness of an initial value problem. In a well-posed initial value problem, the solution of the problem exists, is unique, and depends continuously on the given initial data. No particular initial data are assumed here, and only the first condition is imposed, which means some solutions of the problem should exist (for at least some of the possible initial data). Why has so weak a notion been postulated? Initial value problems that possess non-unique solutions can be interpreted as describing indeterministic evolution. Even if there are reasons to exclude this type of evolution, it should be done on different grounds. The bare notion of spacetime (or, in the present consideration, a basic object of GR) is unrelated to the issue of whether the physical world is deterministic or indeterministic, so the latter should not be settled at the level of specifying what spacetime/a basic object of GR is.

Regarding the second issue (i.e., the benefits of the above sketched functional approach), I argue [7] that the other functions of spacetime one can think of (e.g.,

⁵ Technically, this means that only globally hyperbolic manifolds allow foliations by Cauchy surfaces (see, e.g., [8, p. 395]).

providing the localization for matter, the causal structure, and the structure of symmetries) are best understood as derivative with respect to the above explained main function or as a part of it. Furthermore, I argue that within this approach, one can easily justify why some models of spacetime should be regarded as excessive (namely, those that contain more structure than is needed to support the laws of the evolution-distribution of matter) and, relatedly, why so-called symmetry-to-reality inferences can be performed for spatiotemporal quantities.

Functional considerations do not put sufficient constraints on the theory, as the proposed role can be played by different mathematical structures. GR is clearly not the only known spacetime theory (and the other important theories include Classical Mechanics and Special Relativity); in fact, most physical theories involve some ingredient that corresponds to spacetime. Therefore, some further specification is required, and in the case of GR, this specification consists of choosing a differential geometry framework and the basic scheme of the model of GR, namely $\langle W, g_{\mu\nu} \rangle$ (where W is a differential manifold and $g_{\mu\nu}$ is a metric).

Which of the stages should be regarded as defining what a basic object of GR is? This is to some extent a matter of terminological choice, but I suggest that the natural answer is that the first two stages, namely kinematic and dynamical, specify what is it to be a basic object of GR. The third stage is better understood as choosing from among the different basic objects of GR those that are in agreement with some of the known general features of the actual physical world (such as the general properties of causal structure, in the case of causality conditions) rather than further specifying the notion of a basic object of GR. This seems to be in line with the usual way of speaking: a badly behaving GR spacetime (i.e., one violating some of the physical reasonability conditions) is still counted as a GR spacetime; and a spacetime that is not a solution of Einstein's equations is not a GR spacetime (but it can be a spacetime of some other theory).

5 Generalised Manifolds

Is the above definition of a basic object of GR (given by the requirements of kinematic and dynamical stage) in exact accordance with physicists' practice? Typically, the definition of a differential manifold, which represents a basic object of GR, includes the so-called Hausdorff condition:

Definition 1 (*Hausdorff condition*) For any two points $p, q \in W$, there exist open⁶ sets O_p, O_q such that $p \in O_p, q \in O_q$, and $O_p \cap O_q = \emptyset$.

The condition says that any two points can be separated by open sets in the sense that we can always find two open disjoint sets, one of which includes the first point, and the other includes the second point. This should be distinguished from the local Euclidicity condition, according to which the neighbourhood of every point can be mapped into a subset of \mathbb{R}^n for some n , which is the same for all points of the manifold. The latter is a defining condition of the manifold and is needed for the basic notions

⁶ Open in the manifold's topology, defined by its atlas.

of differential geometry to be definable; it implies the local version of the Hausdorff condition⁷ but not the standard, global one as defined above.

One can consider generalised manifolds, which do not need to satisfy the Hausdorff condition. Then, Hausdorff manifolds (satisfying the condition) and non-Hausdorff manifolds (violating the condition) become two special cases of generalised manifolds. This terminology will be used in what follows.

Although the Hausdorff condition seems like an abstract, purely technical constraint on mathematical formalism, in the framework of GR, its violation has consequences of clear physical significance. Some of these consequences led physicists to restrict their attention to manifolds that satisfy this condition. In Sects. 6–8, I will first review properties of non-Hausdorff manifolds and then reflect upon the issue of whether they indeed should be excluded, and if the answer is ‘yes’, at what conceptual stage this should be done.

6 Properties of Non-Hausdorff Manifolds

This section reviews the results concerning properties of differential manifolds that are connected with the Hausdorff condition, most of which are equivalent with this condition. Some of these properties seem to be purely mathematical without a clear physical interpretation, whereas some others have far-reaching physical consequences. Let us begin with the first group of properties.

The Hausdorff condition is a topological condition, so it can be considered in the purely topological context without taking into account differential structure. Let us restrict our considerations to the topological context for a moment. Denote topological manifolds⁸ by X and Y . In Bourbaki’s [9], Munkres’s [10], and Willard’s [11] work, we can find the following mathematical properties that are equivalent to the Hausdorff condition:

- Every net in X (a subset of X indexed by elements of directed set) has a unique limit point.
- Every filter in X has a unique limit point.
- The diagonal $\Delta = \{(x, x) : x \in X\}$ is closed in $X \times X$.
- The intersection of the closed neighbourhoods of any point of X consists of that point alone.

There is also a group of conditions that are not equivalent to the Hausdorff condition but are absent in some non-Hausdorff manifolds and satisfied in all Hausdorff manifolds:

⁷ The local version of the Hausdorff condition requires that every point has a neighbourhood that, considered as a submanifold, satisfies the Hausdorff condition (i.e., a manifold W is said to be locally Hausdorff iff for every point $p \in W$, there exists a neighbourhood of p which satisfies the Hausdorff condition). The local version does not imply the global one; so, if we find a non-Hausdorff pair (i.e., a pair of points that violate the Hausdorff condition) in a locally but not globally Hausdorff manifold, then for each member of this pair, we can find a neighbourhood that does not include the other member of this pair and, as a whole, satisfies the Hausdorff condition.

⁸ A topological manifold is a topological space X for which there exists $n \in \mathbb{N}$ such that for every point $p \in X$, there exists a neighbourhood of p that is homeomorphic with \mathbb{R}^n .

- Any continuous map of a topological space X into a Hausdorff space Y is uniquely determined by its values at all points of a dense subset of X .
- If X is a Hausdorff space, then a sequence of points of X converges to at most one point in X .
- Every compact subset of a Hausdorff space is closed.

The significance of these results for the physical viability of non-Hausdorff manifolds is not clear. For example, the fact that in Hausdorff manifolds, a continuous map is uniquely determined by its values in a dense subset is very useful for mathematicians as it simplifies proofs of some theorems, but, as far as the author knows, there is no physical motivation for this property. More problematic is the lack of a unique limit point of a net or of a sequence as it may be related to bifurcating curves. However, the non-uniqueness of limit points and the existence of such curves are inequivalent issues, and the latter will be analysed separately.

We have seen consequences of the Hausdorff condition for the topology of a differential manifold. What about the differential structure? It turns out that the Hausdorff condition is not needed for defining all the basic concepts of local differential geometry. Hicks' textbook [12] shows how all these concepts can be defined and analysed without assuming this condition. This observation is crucial as these are precisely the notions that are needed to pose the problem of the evolution-distribution of matter, which suggests that there is no reason for excluding non-Hausdorff manifolds already at the kinematic stage.

However, the Hausdorff condition does make some difference for differential manifolds, in addition to the topological properties mentioned above. Adding this condition enables some further constructions, such as a distance function. To see this connection, let us first define (following [12, p. 70]) the notions of pseudo-distance and distance:

Definition 2 A pseudo-metric (i.e., the pseudo-distance between points p and q) is given by⁹

$$d(p, q) := \inf \{ |\gamma| : \gamma \text{ is a smooth curve from } p \text{ to } q \}, \quad (1)$$

where

$$|\gamma| := \int_{s_1}^{s_2} \left(g_{ab} \xi^a \xi^b \right)^{\frac{1}{2}} ds \quad (2)$$

and ξ^a is the tangent field to γ .

Definition 3 A metric (i.e., the distance between points p and q) is a pseudo-metric that satisfies an additional condition: for every $p, q \in W$,

$$d(p, q) = 0 \Rightarrow p = q. \quad (3)$$

If a pseudo-metric is also a metric, then it enables one to distinguish different points. If it is not, then it is allowed that two distinct points are at the distance 0 from each other. The connection between the Hausdorff condition and metric is given by the following theorem [12, p. 71]:

⁹ This is not a general definition of a pseudo-metric, which is axiomatic, but a particular construction in GR satisfying this definition.

Theorem 1 *A connected Riemannian differential manifold \mathcal{W} is Hausdorff iff the pseudo-metric d is a metric.*

It should be stressed that the term ‘metric’ is used here in the topological sense; a metric in this sense is not the same as the metric tensor. The latter is clearly an indispensable element of any model of GR; the former is not necessary, although it may be useful in some constructions. A metric tensor, in contrast to metric in the topological sense, does not require the Hausdorff condition to be definable; in fact, construction of concrete examples of non-Hausdorff manifolds amounts to finding their metric tensors (see, e.g., [13]).

Next, the results of clearer physical importance will be addressed. There are three such results: concerning bifurcating curves, concerning the existence of the maximal extension of a differential manifold, and concerning the relationship between the Hausdorff condition and strong causality. To understand the first result, we first need to define two kinds of bifurcating curves (see [14]):

Definition 4 (*Bifurcate curve of the first kind*) A bifurcate curve of the first kind is a pair of curves C, C' in a manifold W , with $C, C' : [0, 1] \rightarrow W$, such that $C = C'$ on $[0, g]$ and $C \neq C'$ on $(g, 1]$ for some $g \in (0, 1)$.

Definition 5 (*Bifurcate curve of the second kind*) A bifurcate curve of the second kind is a pair of curves C, C' in a manifold W , with $C, C' : [0, 1] \rightarrow W$, such that $C = C'$ on $[0, g)$ and $C \neq C'$ on $[g, 1]$ for some $g \in (0, 1)$.

The distinction between bifurcating curves of the first and second kind is important, because of the following difference between them. Bifurcating curves of the first kind appear in all differential manifolds and are regarded as unproblematic. In contrast, bifurcating curves of the second kind appear only in very special manifolds (in a subclass of non-Hausdorff manifolds), and some of them are bifurcating geodesics. If one interprets geodesics as potential trajectories of free test particles (which is a standard approach), then it is not clear how to think about the situation of particle passing a bifurcation point. One natural interpretation is that such a point indicates an indeterministic event: a particle needs to ‘choose’ one of the branches of a bifurcating curve, and there is nothing in the theory that determines which branch it should take.

There are no such problems in Hausdorff manifolds, which do not allow bifurcating curves of the second kind. However, even some of the non-Hausdorff manifolds do not admit such curves. This issue is related to the way non-Hausdorff manifolds are constructed from Hausdorff ones. This can be done with the use of a so-called gluing map (see [14,15]):

Definition 6 (*Gluing map*) Let $\mathcal{W}_1 = \langle W_1, g_1 \rangle$ and $\mathcal{W}_2 = \langle W_2, g_2 \rangle$ be differential manifolds. Then $\phi : U_{12} \rightarrow U_{21}$, where $U_{12} \subseteq W_1, U_{21} \subseteq W_2$, is a gluing map if

- U_{12} is open, and
- ϕ is an isometry.

One can distinguish two classes of gluing maps: those that are continuously extendible and those that are not, where continuous extendibility is defined in the following way:

Definition 7 (*Continuously extendible gluing*) A gluing map $\phi : A \rightarrow B$, $A \subseteq W_1$, $B \subseteq W_2$ is continuously extendible iff there exist A', B', ϕ' such that $A \subsetneq A' \subseteq W_1$, $B \subsetneq B' \subseteq W_2$, $\phi' : A' \rightarrow B'$, ϕ' is continuous and $\phi'|_A = \phi$.

Then, the theorem relating the existence of bifurcating curves of the second kind and the choice of gluing is as follows [14]:

Theorem 2 *The necessary and sufficient condition for a differential manifold constructed by gluing together Hausdorff differential manifolds to admit bifurcating curves of the second kind is that the gluing be continuously extendible.*

This may look like a theorem about a very special class of non-Hausdorff manifolds, namely those constructed from Hausdorff manifolds with the use of the gluing technique. However, it may be proven [15] that any Hausdorff manifold can be constructed this way, so the theorem turns out to be a universal criterion for distinguishing non-Hausdorff manifolds admitting bifurcating curves of the second kind from those that do not admit them.

The second physically significant property concerns the existence of a maximal extension of a given differential manifold. Extension and maximal extension are defined as follows:

Definition 8 (*Extension, maximal extension*) An extension of a differential manifold $\langle W, g_{ab} \rangle$ is a differential manifold $\langle W', g'_{ab} \rangle$ such that there exists an isometric embedding $\phi : W \rightarrow W'$ satisfying $\phi(W) \subsetneq W'$. An extension $\langle W', g'_{ab} \rangle$ is a maximal extension if it does not admit any further extension.

Clarke's [16] theorem states the relationship between satisfying the Hausdorff condition by a manifold and the existence of a maximal extension of that manifold:

Theorem 3 *Any spacetime that is either Hausdorff or non-Hausdorff without bifurcating curves of the second kind has a maximal extension.*

The lack of a maximal extension is sometimes regarded as problematic (and we will come back to this issue later). For now, let us observe that, again, the Hausdorff condition is not enough to get the problematic result, but an additional factor is needed, and it is the same as in the case of the existence of bifurcating curves of the second kind, namely the fact that the gluing (by which the manifold can be constructed) is continuously extendible.

The third relevant theorem for our considerations concerns causal properties of non-Hausdorff manifolds. Some of them violate so-called strong causality defined as follows [17, p. 134]:

Definition 9 (*Strong causality*) A differential manifold is said to be strongly causal if for all points $p \in W$, and all sufficiently small open sets O containing p , no smooth future-directed timelike curve that starts in O , and leaves O , ever returns to O .

Concerning the relationship between strong causality and the Hausdorff condition, [16] proves the following:

Theorem 4 *A non-Hausdorff spacetime is either not strongly causal or admits bifurcate curves of the second kind.*

If all of the analysed properties (i.e., bifurcating curves of the second kind, the lack of a maximal extension, and the violation of strong causality) are really problematic, then non-Hausdorff manifolds should be regarded as problematic as well because each of them has at least one of these properties. If we restrict to non-Hausdorff manifolds that are constructible with the use of gluing that is not continuously extendible, then we face the problem of the violation of strong causality. However, if we restrict to non-Hausdorff manifolds that are constructible with the use of continuously extendible gluing, then we can perhaps save strong causality (although Clark’s theorem does not guarantee that), but we have to accept bifurcating curves of the second kind (including bifurcating geodesics) and the lack of a maximal extension. Thus, the crucial question is whether these properties are indeed unacceptable, and if so, at which stage (among those presented in Sect. 3) should they be taken into account.

7 Objections to the Use of Non-Hausdorff Manifolds in GR

The aim of this section is to review the objections to non-Hausdorff manifolds considered in the literature, which typically lead authors to disregard such objects altogether. It should be noted that in most of the literature, non-Hausdorff manifolds are omitted without analysing the reasons for such a move in detail; supposedly, the implicit reason is always one of those presented below. All of the objections refer to the properties of non-Hausdorff manifolds analysed in the previous section. The objections are as follows.

(i) First, there are some purely mathematical objections, referring to the failure in non-Hausdorff manifolds of some mathematical theorems premised on the Hausdorff condition. Objections of this type can be found in [18, pp. 199–200]:

The assumption of Hausdorffness is explicitly invoked only sporadically in textbooks on general relativity. But it is implicitly assumed in so many standard results in GTR that dropping it would require a major rewriting of textbooks. Here are two examples of widely used results that depend on Hausdorffness.

- (i) A compact set of a topological space is closed—if the space is Hausdorff.
- (ii) If a sequence of points of a topological space converges, the limit point is unique—if the space is Hausdorff.

(ii) The main argument against non-Hausdorff manifolds refers to the presence of bifurcating geodesics (see Theorem 2). Since geodesics are standardly assumed to be the (potential) worldliness of free test particles, the bifurcation amounts to indeterminism. Some authors claim indeterminism is problematic, as it leads to an unacceptable kind of arbitrariness. For example, Earman [18, p. 200] asks rhetorically, “But how would such a particle know which branch of a bifurcating geodesic to follow?”, suggesting that there is no good answer to this question, but that there should be one. Also, Hájíček disregards non-Hausdorff manifolds for a similar reason, namely, because they break of the “classical causality conception coinciding with determinism” [19, p. 79].

(iii) An objection similar to (ii) concerns whole spacetimes: if we allow non-Hausdorffness, then one may attach non-Hausdorffly additional branches at some

moment in time to each spacetime, and nothing determines which possible branch is realised. Again, quoting Earman [18, p. 202]:

If non-Hausdorff spacetime branching is to be taken seriously, what is needed is a physical theory that prescribes the dynamics of branching. What the advocates of individual branching offer is not a theory but the nostrum that branching takes place when an indeterministic process produces an outcome plus *ad hoc* rules, e.g. there is branching to the future but not towards the past. (...) “outcome of an indeterministic process” is not a term of any theory of physics but something that needs analysis.

(iv) The fact that some non-Hausdorff differential manifolds do not have maximal extensions, as Theorem 3 states, can also be regarded as problematic. Earman [20, pp. 32–33] claims that there are metaphysical reasons to believe that the actual spacetime should be maximal:

Metaphysical considerations suggest that to be a serious candidate for describing actuality, a spacetime should be maximal. For example, for the Creative Force to actualize a proper subpart of a larger spacetime would seem to be a violation of Leibniz’s principles of sufficient reason and plenitude. If one adopts the image of spacetime as being generated or built up as time passes then the dynamical version of the principle of sufficient reason would ask why the Creative Force would stop building if it is possible to continue. However, this image does not sit well with the four-dimensional way of thinking, and in any case it runs into trouble in its own terms: since extensions of spacetime are generally non-unique there may be many ways to continue building and the Creative Force may be stymied by a Buridan’s ass choice (see Clarke [21], pp. 8, 9). Some readers may be shocked by the introduction of metaphysical considerations in the hardest of the “hard sciences.” But in fact leading workers in relativistic gravitation, though they don’t invoke the name of Leibniz, are motivated by such principles (see, for example, [14, p. 262] and [22, p. 253]).

(v) Earman’s [18] other objection concerns the local and global conservation of energy. His argument goes as follows. To properly formulate a local conservation law, an energy-momentum tensor should be continuous and differentiable. However, this entails that when energy ‘travels’ along a bifurcating curve, it has to take both branches because if it only went along one of them, the tensor on another one would be discontinuous. However, global energy conservation would then be violated.

(vi) Additionally, all arguments in favour of strong causality are simultaneously arguments against those non-Hausdorff manifolds that violate it (see Theorem 4). Most of the literature dealing with causality conditions is not concerned with non-Hausdorff manifolds, as the conditions themselves are of great significance in many other contexts in GR. Besides strong causality, many more causality conditions are considered, and the most important one is the chronology condition, which states that no smooth, closed timelike curves exist. Its importance arises from the fact that the violation of this condition raises serious interpretative problems, such as the grandfather paradox. It seems that causal loops would enable one to intervene in the past, and the possibility

of changing the past seems to be inconsistent or at least requires very sophisticated interpretation, for which our causal intuitions are greatly insufficient.

The chronology condition forbids a smooth timelike curve from coming back to the exact same point, whereas strong causality forbids coming back even arbitrarily close to that point. Although the violation of the latter does not seem to lead to causal paradoxes, or if it does, they are at least considerably less obvious, some authors claim that manifolds allowing such a violation should also be disregarded. For example, Hawking [23] says,

it can be seen that, if the strong causality assumption did not hold, W would be on the verge of violating the chronology assumption. That is to say, the slightest disturbance of the metric (by quantum fluctuations, for instance) could lead to the existence of closed timelike curves. It would not seem realistic to suppose that spacetime could judge as accurately as that to avoid a violation of chronology. Thus it would seem physically reasonable to assume strong causality.

(vii) Finally, Penrose [22] criticises non-Hausdorff manifolds on the basis that the only potential application of these objects in physics he regards as viable (in the Everettian interpretation of Quantum Mechanics) has its own internal problems. Let us analyse these objections in more detail. In [22, pp. 592–595] Penrose considers a model in which “the branching takes place along the future light-cones of the points at which ‘observations’ (presumably quantum-mechanical) are made”. According to him, this model

is not altogether implausible. It is possible to envisage, for example, that the branching accompanies a kind of retarded collapse of the wavefunction, where on each branch the wavefunction starts out as a different eigenvector of the operator representing the observation.

This would be, therefore, a possible model of an Everett-type universe. Then, Penrose continues,

Such a model could be viewed as an ‘objective’ description of a world containing some strongly ‘subjective’ elements. One could envisage different conscious observers threading different routes through the myriads of branches (either by chance, say, or perhaps even by the exercise of some ‘free will’).

Penrose criticises such a non-Hausdorff Everettian model. The first problem is that no one has shown that Everettian assumptions in fact lead to such a model. However, this objection is itself not so harmful, as it is also not shown that such a model cannot be constructed; it amounts to saying that this is a potential project rather than a finished theory. Another disadvantage is that this model does not seem to be able to solve the problem of the origin of the direction of time. And Penrose’s main objection concerns interpretative problems associated with the Everettian conception as such:

I feel particularly uncomfortable about my friends having all (presumably) disappeared down different branches of the universe, leaving me with nothing but unconscious zombies to talk to!

To comment on this quote, Penrose seems to assume that new worlds in the Everettian multiverse could be created by measurements performed by a conscious observer (each world corresponds to one possible outcome), whose consciousness, however, stays only in one of these worlds. In the remaining worlds, only his or her unconscious body remains, and we have no guarantee that the consciousness of people we know stayed in the same world as our own.

8 Answers to the Objections and the Modal Interpretation of Non-Hausdorff Manifolds

How convincing are the above objections, and what conclusions should we draw from them? In terms of our four stages of identifying mathematical structures for physics (discussed in Sect. 3), to what stage do they belong? As mentioned in Sect. 6, non-Hausdorff manifolds enable one to define all the notions of differential geometry needed to pose the problem of the evolution-distribution of matter, which is enough to pass the first (kinematic) stage. Although the solution is not unique in some cases (as bifurcating geodesics can occur, leading to non-unique trajectories of free test particles), it does not matter at this stage because only the fact that the problem can be formulated matters. Its more specific properties should be considered at a later (i.e., the third) stage. With regard to the dynamical stage, nothing in the Einstein equations prohibits the occurrence of non-Hausdorff manifolds, so they can also pass this stage if the stress-energy tensor $T_{\mu\nu}$ is appropriately chosen. Therefore, it seems that if non-Hausdorff manifolds were to be excluded, this should be done at the third or fourth stage. In fact, all of the arguments reviewed in Sect. 7 are associated with the third stage. Concerning the fourth stage, very little can be said. Currently, no successful usage of non-Hausdorff manifolds to represent some physical situation is known, and proposing such a model is beyond the scope of this paper.

The aim of this section is to discuss whether the presented objections are strong enough to lead to the judgement that non-Hausdorff manifolds do not pass the third stage. This is particularly difficult, as the criteria here are not clear: there are different ideas of what counts as physical reasonability conditions, and satisfying some of them does not imply that the others are satisfied as well, so they should be considered case by case. Therefore, the following analysis should be regarded as rather provisional.

Ad (i): First, observe that the two properties Earman mentions are not equivalent to the Hausdorff property, but only implied by it, so it is not the case that all non-Hausdorff spacetimes violate them. Second, and more importantly, as was suggested in Sect. 6, it is not clear whether these mathematical facts have a physical significance on their own. If they are important because they are connected with some more physical features of non-Hausdorff manifolds, such as the presence of bifurcating geodesics, then these physical features, and not the mathematical facts mentioned by Earman, form the proper basis of the argument against non-Hausdorff manifolds.

An alternative reading of Earman's quote could be that the crucial thing is not the *particular* mathematical theorems mentioned by him, but a *general* fact (of a rather practical nature) that all the available literature assumes the Hausdorff condition and, therefore, we would be unable to recover the results that are important for scientific

practice if we dropped it. However, this is not the case—although most of the literature on the subject indeed assumes the Hausdorff condition, not all textbooks do this. For example, as previously mentioned, [12] introduces all the basic notions of differential geometry without assuming the Hausdorff condition and starts using it substantially no earlier than in chapter 6 of his book.

Before we move on to the next objection, we should recall the exact relationship between non-Hausdorff manifolds and bifurcating geodesics. As Theorem 2 states, some non-Hausdorff manifolds do *not* admit bifurcating curves of the second kind (some of which may be bifurcating geodesics). Interestingly, most of the non-Hausdorff manifolds considered in the existing literature are of this type, for example, extensions of Misner spacetime and extensions of Taub-NUT spacetime (see [13,14,24,25]). Therefore, some physicists do consider liberalizations of the Hausdorff condition. For example, Hawking [25, p. 174] allow for these non-Hausdorff manifolds that do not admit bifurcating geodesics. Similarly, Geroch [26, p. 465] allows for non-Hausdorff manifolds in which every geodesic has a unique extension and every curve has no more than one end point.

As a consequence, our initial question of whether non-Hausdorff manifolds are physically reasonable splits into two: (1) are non-Hausdorff manifolds with bifurcating geodesics physically reasonable (which is addressed in the objections (ii)–(v)) and (2) are non-Hausdorff manifolds without bifurcating geodesics, which, according to Theorem 4, violate strong causality, physically reasonable (which is addressed in the objection (vi))?

Ad (ii)–(iii): The main problem that arises from the violation of the Hausdorff condition is indeterminism, both at the level of geodesics (i.e., curves followed by free test particles) and at the level of the entire manifold (i.e., non-Hausdorffly attached ‘branches’). However, it is far from obvious that the fact that a theory involves indeterminism should be considered as an argument against it. For this argument to be valid, the authors should provide some independent argument that indeterminism is not allowed in physical theories (which seems to be too strong a position, as physicists in general allow indeterministic theories) or that the specific form of indeterminism that occurs in the case under consideration is unacceptable. Without such further support, this argument comes down to a simple rejection of indeterminism.

One way of supplementing this argument is to say that indeterminism is problematic because it is a kind of arbitrariness and therefore should be avoided in science. The issue of whether some arbitrariness exists in nature (and if so, what types of arbitrariness are allowed) is interesting but somewhat speculative.¹⁰ It seems that in our explanation of physical facts, we need to stop at some point anyway, which means that some arbitrary factor needs to be postulated. It can be either the value of some physical constant or the form of the equation; or, if we managed to derive both equations and constants, the arbitrary factor would be the assumption from which they are derived. This is

¹⁰ The word ‘arbitrariness’ may seem to be ambiguous, as it can mean (1) that something is underdetermined, (2) that something lacks an explanation, or (3) that something is not objective. I assume that all considerations of this paper concern issues that are objective, so I am not interested in the third sense. Furthermore, I treat the first sense as a special case of the second: if the evolution of a physical system is not deterministic, then there is no explanation why this particular trajectory or history has been realised rather than some other that is also allowed by the laws.

arbitrariness at the level of general facts about our world (laws, constants, etc.). What about the level of particular facts (i.e., why this particular course of events have been realised from among the bigger set of courses of events allowed by what we called ‘general facts’)? Can we *a priori* exclude that something unexplainable is already at the level of particular facts as indeterminism postulates? Do we have reasons to insist that all the arbitrariness arises at the general level, that is, at the level of the structures that explain particular facts, such as laws and physical constants? These are difficult questions, but they have been intentionally posed in a way to suggest that there are no obvious answers to them. And such answers, if attainable at all, should not be found *a priori*, but on the basis of considering our best empirical theories; if some of them are best interpreted as indeterministic, we are allowed to ascribe (tentatively) indeterminism to the physical reality itself. The argument from indeterminism was presented by Earman and Hájíček in a way suggesting that the answer is known *a priori* or at least obvious on the basis of the empirical knowledge we already have, but this is clearly not the case.

How should bifurcating geodesics or new branches attached to a manifold be conceptualized? An interesting interpretative option is that a basic object of GR does not necessarily represent a single spacetime. In the case of non-Hausdorff manifolds, another interpretation, which can be called ‘modal’, is available, namely that they represent bundles of alternative possible spacetimes, only one of which is actual. This interpretation can be supported by the following theorem (for the proof, see [15]):

Theorem 5 *Any non-Hausdorff manifold can be constructed by gluing together a family of Hausdorff manifolds.*

The modal interpretation, based on the above theorem, goes as follows. Any non-Hausdorff manifold can be decomposed into a family of Hausdorff submanifolds¹¹ (from which it can be obtained by gluing maps, see definition 6), and each of these Hausdorff manifolds represents a single possible spacetime.¹² A non-Hausdorff manifold represents all the possibilities, taken together, but exactly one of them actualises. Things look similar at the level of geodesics. A bifurcation of geodesics requires non-Hausdorffness to occur, which means that each branch of a geodesic can be associated with a single spacetime, and bifurcating geodesics as a whole can be interpreted as representing the possible trajectories of free test particles, exactly one of which actualises.¹³ This is why the notion of the basic object of GR instead of just spacetime has been used throughout the text; according to the modal interpretation, a single spacetime is not the only object that can be represented by a manifold.

One may worry that such an interpretation leads us too far away from the original theory: if we include the subject matter of a theory in its identity conditions and assume that the subject matter of GR is spacetime, then the above change of the

¹¹ Such a decomposition is not unique; however, if we restrict to maximal Hausdorff submanifolds, it will be unique.

¹² A similar idea in the context of the theory of branching spacetimes has been developed by [27–30]: the models of the branching spacetime axioms are non-Hausdorff, but histories (i.e., constructs representing single spacetimes) are Hausdorff.

¹³ Therefore, we obtain here indeterminism in two senses: global for whole spacetimes and local for geodesics; for more on this distinction, see [31,32].

subject matter would amount to changing the theory into a new one. To avoid the mentioned assumption, I introduced in Sect. 2 the notion of the basic object of GR, allowing in principle that it can be different from spacetime. In Sect. 4, a functional restriction was placed on it, namely that it should play a certain role with respect to matter; and in Sect. 5 it was stated that non-Hausdorff manifolds can play this role. ‘Can play’ means ‘have enough mathematical structure’ to do so; whether they actually do can depend on the physical interpretation, but surely the modal interpretation does not prevent them from doing so. Another way to convince ourselves that this is still the same theory is to define its subject matter more ‘locally’, namely, instead of saying that GR is about spacetime (or about a basic object of GR as defined earlier), one can say that it describes spatiotemporal states of affairs (and every model of GR describes many such states). Whether a given state of affairs is spatiotemporal does not depend on whether it is actual, so merely possible spatiotemporal states of affairs are still subjects of GR; and under the modal interpretation, non-Hausdorff manifolds describe (actual and possible) spatiotemporal states of affairs no less than Hausdorff manifolds, interpreted as single spacetimes, do.

Ad (iv): Earman’s quoted argument against extendible manifolds relies on reluctance to accept arbitrariness, which is the issue addressed while discussing objections (ii)–(iii). However, two further points are to be made concerning this particular objection. First, some of the formal results suggest that in some cases, extendible manifolds should be preferable over inextendible ones, as is summarized by Manchak [5, pp. 415–416], who says:

But, however compelling the metaphysics, it is sometimes problematic to insist on inextendibility. For example, [16, p. 20] has shown that not every well-behaved space-time admits a well-behaved inextendible extension. Should we cling to inextendibility at the expense of other desirable space-time properties? The answer is far from clear. Additionally, a space-time does not always have just one inextendible extension [21, p. 9]. Thus, the principle of sufficient reason can actually be used to argue against the property of inextendibility. After all, why should one extension be preferred over another?

Therefore, even if we accept the type of reasoning based on the principle of sufficient reason, it is not clear how it should work in the case of extendible manifolds. In some cases, it is more important not to lose some other properties that our model of spacetime is expected to have than to save inextendibility; and in some other cases, more than one extension is available, and the choice between them is arbitrary, unless one of them is distinguished by special properties the others lack.

The second point I want to make is specific to the context of the modal interpretation of non-Hausdorff manifolds. Adding new branches (at the level of mathematical construction) to the already constructed manifold does not amount to enlarging the model of the actual spacetime, as only one of these branches, at most, is actualised under the modal interpretation. Therefore, what is extended is the structure of possible alternatives for a spacetime to evolve, not the spacetime itself. To what extent does this alleviate Earman’s worry? This worry can be formulated in two ways: one in terms of the ‘Creative Force’ and the other in terms of the existence of sufficient reason,

but both formulations concern only the actualised part of the manifold; namely, they state in different ways the question, ‘why has this particular spacetime been actualised and not a larger one?’ Therefore, they ask about a single spacetime (the one that is actualised), not about the whole bunch of spacetimes. And for the actualised part, there is a good answer why it should not be extended to the whole non-Hausdorff structure: because we identified a single spacetime with a Hausdorff submanifold.

Ad (v): Another objection posed by Earman, concerning energy, turns out to be harmless if we carefully interpret branching structures as representing possible evolutions, where one branch at most can be actualised. In the actual world, the energy tensor is wholly contained in one branch, and the discontinuity concerns only branches that are not realised. Therefore, if the continuity condition is violated at all, it is not violated by any actual physical object, only by non-actual ones. What about a counterfactual situation in which a different branch had been actualised? Then energy would follow that other branch which is actualised in that counterfactual situation. There would never be a problem of discontinuity in the actual physical reality. As we do not need to postulate that energy follows all the branches, the problem with the conservation of energy does not arise.¹⁴

Ad (vi): As already mentioned in the exposition of this objection, it is not the violation of strong causality that is problematic, but the violation of a weaker condition, called the chronology condition; strong causality is perceived as a desired feature because of its close relation with the chronology condition. According to the quote by Hawking, we should disregard manifolds that are not strongly causal because they are, in some sense, very close to manifolds that contain closed timelike loops.

Under which assumptions should this closeness be regarded as problematic? Hawking appeals to the idea of ‘nature judging accurately’, which seems to be a kind of fine-tuning, and to quantum fluctuations. We have to deal with fine-tuning if, in a space of possible models, the model that is supposed to be the actual one is very close to other models that are radically different, where ‘closeness’ can be measured, for example, in terms of values of some physical parameter with respect to which the models differ. This situation is regarded as problematic because it makes some important characteristics of the actual model improbable; to get the actual model, the values of the appropriate parameters should be ‘fine-tuned’. This type of reasoning is often invoked in the context of particle physics.¹⁵ In our context, it is not clear which parameter would be fine-tuned, but it is clear that almost-closed timelike curves are, in some sense, ‘close’ to closed timelike curves.

A typical objection to fine-tuning arguments is that they require an assumption about a probability measure on the space of possible models, and there is no way to infer such a measure from what is observed. Furthermore, if the choice of parameters is a single, unrepeatable ‘event’, then the whole idea of describing it probabilistically is suspicious. Perhaps the origin of a probability measure, in our context, should be from quantum theory as the invoking of quantum fluctuations by Hawking suggests. If we allowed for manifolds violating strong causality—the reasoning can go—then

¹⁴ Another possible way of argumentation is that in GR the notion of global energy conservation is problematic anyway.

¹⁵ For a review of the issue of fine-tuning in general, see [33]. It is investigated critically in the context of particle physics, for example, by [34–36].

we would also allow for all the manifolds that are obtainable from them by quantum fluctuations, and manifolds that violate chronology are probably among them. But this ‘probably’ is supported by intuition rather than calculation; one cannot exclude that quantum laws prevent somehow quantum fluctuations from entering the chronology-violating region (in the space of models) even if one starts from manifolds violating strong causality.

This analysis is clearly inconclusive, but at least two messages can be drawn from it. First, the violation of strong causality alone does not lead to the interpretative problems connected with causal loops. Second, starting from the chronology condition, we need some quite substantial additional assumptions to justify strong causality as a physical reasonability condition.

Ad (vii): Penrose’s arguments do not concern non-Hausdorff manifolds as such, but only one possible way of using them in physics. The first argument concerns the fact that no detailed model connecting non-Hausdorff manifolds with the known forms of indeterminism has been built, but the possibility of such a model is not excluded by Penrose. It is also not obvious that the only motivation for using non-Hausdorff manifolds can come from the Everettian interpretation of Quantum Mechanics, as Penrose seems to implicitly assume; other indeterministic interpretations of Quantum Mechanics can provide an equally good motivation, and of course, one does not need to restrict to Quantum Mechanics (although it is currently the most serious candidate for a theory capturing genuinely indeterministic phenomena).

Regarding the second argument, solving the issue of the direction of time does not need to be a role of non-Hausdorff manifolds. The question under investigation is whether they may form a basis for physical models of the actual world (or at least of some physically possible world), not what further problems can be solved by using them. The idea that some relationship between indeterminism and the direction of time exists is attractive because indeterminism is a temporally asymmetric phenomenon: there is only one way for the past to be (namely, the actual one), but there is more than one way for the future to be (as it is not unambiguously determined). However, it is not yet clear whether the two asymmetries are indeed the same or even deeply related, and more popular discussions of the direction of time do not refer to indeterminism at all (see, e.g., [37, part II]). However, attempts have been made to try to relate the asymmetry of past and future with the asymmetry involved in indeterminism (see, e.g., [38,39]), so if they are accurate and if the correct representation of indeterminism involves non-Hausdorff manifolds, Penrose’s claim may turn out to be false.

Penrose’s third objection can be decomposed into three steps: (1) the only motivation for using non-Hausdorff manifolds in physics is the Everettian interpretation of quantum mechanics; (2) this interpretation commits us to paradoxical theses about minds and their relationship to the world; from (1) and (2), it follows that (3) non-Hausdorff manifolds have no place in physics. We have already seen that the first premise is highly contestable. And the second premise is clearly false: the Everettian interpretation should not be read as assuming anything about the nature of mind and consciousness—it is not the same as the ‘many-minds’ interpretation. Instead, it should be called the many-worlds interpretation or the emergent multiverse interpretation, as in Wallace’s [40] book. Therefore, the argument is more about potential problems for the many-minds interpretation of Quantum Mechanics and does not threaten non-

Hausdorff manifolds as such, as using models based on non-Hausdorff manifolds does not commit us in any way to any hypotheses about the relationship between the physical world and our minds.

9 Two Senses of Modality

In considerations concerning the use of non-Hausdorff manifolds as basic objects of GR, two senses of modality were engaged. The first one, connected with indeterminism, was the basis of the modal interpretation of non-Hausdorff manifolds as representing a bundle of alternative possible spacetimes rather than a single spacetime. The second one is connected with the different stages of determining the proper structure to describe the physical world: the object that passes more stages can be regarded as ‘more possible’ or possible in a stronger sense, and the whole path culminates in reaching the mathematical structure that represents the actual physical world.¹⁶

In the case of GR, and perhaps other theories, some of the dynamically possible models have properties regarded as pathological, so they are excluded as physically unreasonable. One can propose that this exclusion corresponds to the difference in the level/strength of possibility: physically reasonable models are ‘more’ possible than merely dynamically possible ones, which are ‘more’ possible than those that are merely kinematically possible, but only one of them is the model of the actual physical world. This proposal assumes that it makes sense to talk about the levels or strengths of modality, which amounts to two claims: that there are many types of modalities and that they can be graded linearly or partially according to some criteria. Concerning the first claim, it has been defended recently by some metaphysicians (e.g., [41–43]). Concerning the second claim, at least partial grading is typically assumed in theories that allow many types of modality; a more familiar example could be the distinction between mathematical or logical and physical possibility, the first of which is weaker (and the respective necessity is stronger). The picture of the types of modalities associated with the first three stages of Sect. 3 would amount to adding a similar grading within physical possibility and necessity itself.

Are these different levels of physical modality a genuine metaphysical feature of the world or merely the reflection of the way we think about it and construct models? Perhaps it only captures the fact that some features of the world are more easily identifiable than others, but in reality, they are all on a par. However, the other option is that there is indeed some hierarchy of importance between them, which is mirrored by different ‘strengths’ of modality associated with different stages of constructing a theory.

To make the picture even more complicated, the hierarchy of the four stages suggested here can be contested. For example, Curiel [6] suggests that energy conditions (not discussed in this paper), which belong to physical reasonability conditions, are in some sense, more fundamental than other physical reasonability conditions and even than the Einstein equations because they are applicable across a wide range of theories of spacetime (not only in GR), and we do not know how to prove them starting from

¹⁶ ‘Actual’ with respect to the modality of the second type, so it can be itself a modal structure in the first, indeterministic sense.

some other conditions¹⁷ (while in contrast, many other physical reasonability conditions admit such proofs). Nevertheless, to formulate them mathematically, we need a general framework, so we need to go at least through the kinematic stage. Additionally, the typical use of such conditions is to narrow down the set of dynamically possible models, so in the exposition of a physical theory, their place is after both the kinematic and the dynamical stage.

10 Summary

From the above considerations, at least two conclusions emerge. First, non-Hausdorff manifolds can be regarded as basic objects of GR, along with Hausdorff manifolds. If they were to be excluded, this would be done at the level of the physical reasonability stage – but even this is not clear, as argued in Sect. 8. I do not postulate that scientific practices or the way in which textbooks are written should change, but the message of this paper is that we should bear in mind the conceptual difference between violating one of the defining conditions for a basic object of GR and violating one of the physical reasonability conditions. This observation, however, opens the door for potentially extending the scope of relativistic investigations so that they include non-Hausdorff manifolds in addition to Hausdorff ones.

The second conclusion is that non-Hausdorff manifolds are better interpreted as bundles of alternative possible spacetimes rather than as a single spacetime. The above two conclusions are connected with the two senses of modality discussed in Sect. 9. Determining the exact metaphysical and epistemological significance of these two senses seems to be a topic worth investigating for philosophers of physics.

Acknowledgements I would like to thank Tomasz Placek for many discussions about the topic of this paper; I also thank the audiences of *Thursday Seminars* in Oxford, *British Society for the Philosophy of Science Annual Conference* in Oxford, *Modality in Physics* conference in Kraków and *3rd Logic, Relativity and Beyond Conference* in Budapest, as well as an anonymous referee for many interesting comments. This work is supported by the National Science Centre in Poland (research grant Opus 2016/23/B/HS1/00464).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Allori, V., et al.: On the common structure of Bohmian mechanics and the Ghirardi-Rimini-Weber theory. *Br. J. Philos. Sci.* **59**(3), 353–389 (2008)
2. Allori, V.: Primitive ontology and the structure of fundamental physical theories. In: Albert, D.Z., Ney, A. (eds.) *The Wave Function. Essays on the Metaphysics of Quantum Mechanics*, pp. 58–75. Oxford University Press, Oxford (2013)
3. Belot, G.: The representation of time and change in mechanics. In: Butterfield, J., Earman, J. (eds.) *Philosophy of Physics: Part A*, pp. 133–227. Elsevier, Amsterdam (2007)
4. Healey, R.: Perfect symmetries. *Br. J. Philos. Sci.* **60**(4), 697–720 (2009)

¹⁷ Which he regards as an indication that no natural proof is available.

5. Manchak, J.B.: What is a physically reasonable spacetime? *Philos. Sci.* **78**, 410–420 (2011)
6. Curiel, E.: A primer on energy conditions. In: Lehmkuhl, D., Schiemann, G., Scholz, E. (eds.) *Towards a Theory of Spacetime Theories: Einstein Studies*, vol. 13, pp. 43–104. Birkhäuser, New York (2017)
7. Luc, J.: What does spacetime do? *Manuscript* (2019)
8. Choquet-Bruhat, Y.: *General Relativity and the Einstein Equations*. Oxford University Press, Oxford (2009)
9. Bourbaki, N.: *General Topology*. Reading Massachusetts. Addison-Wesley Publishing Company, Boston (1966)
10. Munkres, J.R.: *Topology*, 2nd edn. Prentice Hall, Upper Saddle River, NJ (2000)
11. Willard, S.: *General Topology*. Addison-Wesley Publishing Company, Reading (1970)
12. Hicks, N.J.: *Notes on Differential Geometry*. Van Nostrand Reinhold Company, New York (1965)
13. Chruściel, P.T., Isenberg, J.: Nonisometric vacuum extensions of vacuum maximal globally hyperbolic spacetimes. *Phys. Rev. D* **48**(4), 1616–1628 (1991)
14. Hájíček, P.: Bifurcate space-time. *J. Math. Phys.* **12**, 157–160 (1971a)
15. Luc, J., Placek, T.: Interpreting non-Hausdorff (generalized) manifolds in General Relativity. *Philos. Sci.* (2019)
16. Clarke, C.J.S.: Space-Time Singularities. *Commun. Math. Phys.* **49**, 17–23 (1976)
17. Malament, D.B.: *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. The University of Chicago Press, Chicago (2012)
18. Earman, J.: Pruning some branches from branching spacetimes. In: Dieks, D. (ed.) *The Ontology of Spacetime II*, pp. 187–206. Elsevier, Amsterdam (2008)
19. Hájíček, P.: Causality in non-Hausdorff space-times. *Commun. Math. Phys.* **21**, 75–84 (1971b)
20. Earman, J.: *Bangs, Crunches, Whimpers, and Shrieks. Singularities and Acausalities in Relativistic Spacetimes*. Oxford University Press, Oxford (1995)
21. Clarke, C.J.S.: *Analysis of Space-Time Singularities*. Cambridge University Press, Cambridge (1993)
22. Penrose, R.: Singularities and time-asymmetry. In: Hawking, S.W., Israel, W. (eds.) *General Relativity: An Einstein Centenary Survey*, pp. 581–638. Cambridge University Press, Cambridge (1979)
23. Hawking, S.W.: *Singularities and the Geometry of Space-Time*, Adams Prize Essay, Department of Applied Mathematics and Theoretical Physics, Cambridge University. Published later in *European Physical Journal H* **39**, 413–503 (2014) (1966)
24. Hájíček, P.: Extensions of the Taub and NUT spaces and extensions of their tangent bundles. *Commun. Math. Phys.* **17**, 109–126 (1970)
25. Hawking, S.W., Ellis, G.F.R.: *The Large Scale Structure of Space-Time*. Cambridge University Press, Cambridge (1973)
26. Geroch, R.: Local characterization of singularities in general relativity. *J. Math. Phys.* **9**, 4504–65 (1968)
27. Placek, T., Belnap, N.: Indeterminism is a modal notion: branching spacetimes and Earman’s pruning. *Synthese* **187**(2), 441–469 (2010)
28. Placek, T., Belnap, N., Kishida, K.: On topological aspects of indeterminism. *Erkenntnis* **79**, 403–436 (2014)
29. Müller, T.: A generalized manifold topology for branching space-times. *Philos. Sci.* **80**(5), 1089–1100 (2013)
30. Placek, T.: Branching for general relativists. In: Müller, T. (ed.) *Nuel Belnap on Indeterminism and Free Action*, pp. 191–221. Springer, New York (2014)
31. Placek, T.: Laplace’s demon tries on Aristotle’s cloak: on two approaches to determinism. *Synthese* **196**(1), 11–30 (2019)
32. Müller, T., Placek, T.: Defining determinism. *Br. J. Philos. Sci.* **69**(1), 215–252 (2018)
33. Friederich, S.: Fine-tuning. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition). <https://plato.stanford.edu/archives/win2018/entries/fine-tuning/> (2018)
34. Grinbaum, A.: Which fine-tuning arguments are fine? *Found. Phys.* **42**, 615–631 (2012)
35. Williams, P.: Naturalness, the autonomy of scales, and the 125 GeV Higgs. *Stud. Hist. Philos. Mod. Phys.* **51**, 82–96 (2015)
36. Hossenfelder, S.: *Lost in Math: How Beauty Leads Physics Astray*. Basic Books, New York (2018)
37. Callender, C. (ed.): *The Oxford Handbook of Philosophy of Time*. Oxford University Press, Oxford (2011)
38. Placek, T.: A locus for “now”. In: Dieks, D., et al. (eds.) *Explanation, Prediction, and Confirmation*, pp. 395–410. Springer, Dordrecht (2011)

39. Placek, T.: Past, present and future modally introduced. *Synthese* (2019). <https://doi.org/10.1007/s11229-019-02302-w>
40. Wallace, D.: *The Emergent Multiverse*. Oxford University Press, Oxford (2012)
41. Fine, K.: Varieties of necessity. In: Gendler, T.S., Hawthorne, J. (eds.) *Conceivability and Possibility*, pp. 253–281. Oxford University Press, Oxford (2002)
42. Müller, T.: Branching in the landscape of possibilities. *Synthese* **188**(1), 41–65 (2012)
43. Williamson, T.: Modal science. *Can. J. Philos.* **46**(4–5), 453–492 (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.